

# EXHIBIT 4

**MORRISON FOERSTER**

425 MARKET STREET  
SAN FRANCISCO  
CALIFORNIA 94105-2482  
  
TELEPHONE: 415.268.7000  
FACSIMILE: 415.268.7522  
  
WWW.MOFO.COM

MORRISON & FOERSTER LLP  
AUSTIN, BEIJING, BERLIN, BOSTON,  
BRUSSELS, DENVER, HONG KONG,  
LONDON, LOS ANGELES, MIAMI,  
NEW YORK, PALO ALTO, SAN DIEGO,  
SAN FRANCISCO, SHANGHAI, SINGAPORE,  
TOKYO, WASHINGTON, D.C.

Writer's Direct Contact  
+1 (213) 892-5200  
ABennett@mofocom

November 6, 2023

***Via Email***

Joseph Saveri  
JOSEPH SAVERI LAW FIRM, LLP  
601 California Street, Suite 1000  
San Francisco, California 94108  
jsaveri@saverilawfirm.com

Re: *Doe 1 v. GitHub, Inc.*, No. 4:22-cv-06823-JST (N.D. Cal.) & *Doe 3 v. GitHub, Inc.*,  
No. 4:22-cv-07074-JST (N.D. Cal.) OpenAI's Responses to Plaintiffs' First Set of  
Interrogatories and Requests for Production

Dear Joseph:

We write in response to your September 20, 2023 letter to clarify OpenAI's positions on its responses to Plaintiffs' First Set of Interrogatories and Requests for Production, following the parties' September 5 and September 7, 2023 meet and confers. At this juncture, OpenAI cannot assess whether the parties are at an impasse because, as detailed further below, OpenAI is awaiting additional information from Plaintiffs.

OpenAI reserves the right to supplement its positions as to the requests addressed herein, as well as Plaintiffs' other discovery requests not addressed in this letter.

**Interrogatories**

**Verifications.** OpenAI will provide verifications to its responses to Plaintiffs' First Set of Interrogatories as soon as reasonably practicable (and before depositions of OpenAI defendant witnesses).

**Interrogatories Nos. 1 and 14.** OpenAI will supplement its responses to identify Bates numbers of responsive documents, to the extent they exist and pursuant to OpenAI's

## |||HARRISON FOERSTER

Joseph Saveri  
November 6, 2023  
Page Two

objections and responses, as soon as reasonably practicable. To the extent information is made available through other means pursuant to the ESI order entered in this case (e.g., made available for inspection on a secure source code computer pursuant to the Amended Protective Order or through querying or sampling methodologies), OpenAI's supplemental responses will so state.

**Interrogatory No. 2.** OpenAI objected to this interrogatory on multiple bases, including that Plaintiffs have failed to explain the relevance of the names of “*all* past and present directors and officers” (emphasis added). Indeed, even if certain officers or directors have limited knowledge of codes, that would not necessarily make their names relevant. Please provide an explanation for why this interrogatory seeks discovery that is both relevant and proportional to the needs of the case so that we can properly assess Plaintiffs' position and whether the parties are at an impasse.

**Interrogatory No. 3.** As stated in our response and during the parties' meet and confer, “all individuals or entities who possess or have possessed stock or ownership interest in [OpenAI] greater than five percent” is irrelevant to any issue in this litigation. Plaintiffs still have not provided an explanation for why an identification of all such individuals is warranted. During the meet and confer, Plaintiffs indicated that this interrogatory seeks to identify the individuals and entities that control Defendants, but as OpenAI expressed during the meet and confer, that is not what this interrogatory seeks as worded. OpenAI stated that it would consider responding to a revised interrogatory specifically seeking this information. Plaintiffs have not provided a revised interrogatory.

**Interrogatory No. 4.** OpenAI disagrees with Plaintiffs' characterization that OpenAI agreed to “answer for the agreements identified in response to Interrogatory 1.” OpenAI did not—and does not—agree to produce information with respect to the individuals who negotiated those agreements. Negotiations for agreements with Microsoft are not relevant to any of the claims asserted here. Plaintiffs still have not explained how agreements or negotiations with Microsoft unrelated to Copilot and/or Codex bear any relevance to the claims at issue. As OpenAI indicated during the parties' meet and confer, OpenAI is considering and further assessing whether OpenAI would be willing to identify the individual(s) responsible for negotiating *aspects* of its agreements with Microsoft specific to Copilot and/or Codex. To the extent OpenAI determines that identifying such individual(s) would be appropriate, OpenAI will supplement its response accordingly.

**Interrogatory No. 5.** Plaintiffs are not entitled to the identities of all persons who have had *any* knowledge about the “engineering and development of Copilot and its previous iterations, versions, or forms” because these witnesses would have knowledge duplicative and/or cumulative of the witnesses already identified in OpenAI's response to this interrogatory. Plaintiffs also have not explained why additional witnesses would be relevant and proportional to the needs of the case here, and OpenAI will not identify every single

## **HORRISON FOERSTER**

Joseph Saveri  
November 6, 2023  
Page Three

OpenAI employee who has any iota of knowledge related to Codex, no matter how trivial, inconsequential, or cumulative of the identified individuals' knowledge.

**Interrogatories Nos. 6 and 7.** OpenAI disagrees that its objections are not appropriate. It is unclear what you mean in your letter by limiting these interrogatories to "identifying OpenAI-affiliated individuals." OpenAI already explained that these interrogatories seek information that does not exist because OpenAI has never purchased or sold any interest in GitHub. OpenAI is willing revise its responses to confirm this statement if such revisions would resolve any dispute pertaining to these interrogatories.

**Interrogatory No. 8.** Plaintiffs have not explained how information related to GPT-3 and GPT-4 falls within the scope of the allegations in this litigation. Please explain the relevance of GPT-3 and GPT-4, and identify specific paragraphs in the Amended Complaint disclosing the theory under which you claim GPT-3 and GPT-4 are relevant.

**Interrogatory No. 9.** OpenAI states again that it does not have "oversight or control over engineering, developing, marketing, or sale of Copilot." Your attempt to seek information from OpenAI that is outside of its possession, custody, and/or control is inappropriate. To the extent Plaintiffs seek information regarding aspects of Copilot that use the same underlying technology as Codex, OpenAI has agreed to produce relevant information about Codex in response to other interrogatories and requests for production. To the extent Plaintiffs seek information regarding the names of Microsoft and/or GitHub individuals with this knowledge, this interrogatory is better directed to Microsoft and/or GitHub.

**Interrogatory No. 10.** Nowhere in Plaintiffs' interrogatories do Plaintiffs define "Ethics Team," and when asked during the meet and confer where this term was defined in the Amended Complaint or elsewhere, Plaintiffs could not do so. We are therefore "puzzled" by your claim that you are "puzzled" by OpenAI's confusion. With the clarification in Plaintiff's September 20, 2023 letter that this interrogatory seeks information regarding "[individuals] involved with OpenAI's 'Safety & Alignment' projects," OpenAI is willing to investigate and provide an updated response, if appropriate, as soon as reasonably practicable.

**Interrogatory No. 11.** As stated in OpenAI's response and during the meet and confer, this interrogatory specifically asks for information regarding Copilot, which is not an OpenAI product. To the extent Plaintiffs seek information regarding aspects of Copilot that use the same underlying technology as Codex, OpenAI has agreed to produce relevant information about Codex in response to other interrogatories and requests for production. OpenAI has been investigating, and will continue to investigate, third-party products used to maintain documents related to Codex and source code for Codex models, and is willing to supplement its response to this interrogatory as soon as reasonably practicable.

## **|||HARRISON FOERSTER**

Joseph Saveri  
November 6, 2023  
Page Four

**Interrogatory No. 12.** OpenAI disputes that Plaintiffs are entitled to “home addresses, email addresses, social media handles, and telephone numbers” of its former and current employees, especially where these individuals are represented by counsel and OpenAI has stated they can be contacted through its outside counsel of record. Moreover, Plaintiffs have yet to explain why the personally identifying information of these individuals is relevant to this case.

**Interrogatory No. 13.** OpenAI is currently investigating sources of data used to train Codex and will supplement its response as soon as reasonably practicable.

### **Requests for Production of Documents**

**Global Issues.** Plaintiffs raise two global issues in their letter: (1) training as it relates to Copilot, and (2) certain products that fall beyond the scope of this litigation. Regarding the first issue, OpenAI states that it does not exercise any oversight or control over Copilot and that Copilot is not a product offered by OpenAI. To the extent Plaintiffs seek Copilot-specific documents, Plaintiffs should direct these requests to Microsoft and/or GitHub. But to the extent Plaintiffs seek information regarding aspects of Copilot that use the same underlying technology as Codex, OpenAI has agreed to produce relevant information about Codex in response to other requests. Regarding the second issue, OpenAI states that requests seeking documents related to Azure, GPT-3, and GPT-4 do not fall within the scope of the allegations in this case, and thus, are not relevant to this litigation. As requested above with respect to the interrogatories, please explain the relevance of Azure, GPT-3, and GPT-4 and specific paragraphs in the Amended Complaint disclosing that theory of relevance.

**Request For Production No. 1.** As indicated above and made clear during the parties’ meet and confer, Copilot is a product offered by Microsoft and/or GitHub, not OpenAI. To the extent Plaintiffs seek documents sufficient to identify Microsoft and/or GitHub employees responsible for Copilot, such requests are more appropriately directed to Microsoft and/or GitHub.

**Requests For Production Nos. 2, 5-7.** OpenAI disputes that documents related to Azure, GPT-3, and GPT-4 fall within the scope of the allegations in this litigation, as noted above. OpenAI has repeatedly requested that Plaintiffs provide an explanation as to their supposed relevance, and has yet to receive one. OpenAI has therefore agreed to produce only organizational charts that identify individuals primarily responsible for the design and operation of Codex. Regarding Plaintiffs’ request for “data or other native information from which the charts were generated or produced” for “prior periods,” it is unclear what Plaintiffs are referring to. OpenAI is unaware of any data that has been destroyed with respect to organizational charts. Plaintiffs asked during the meet and confer what system OpenAI uses for organizational charts. OpenAI uses ChartHop.

## MORRISON FOERSTER

Joseph Saveri  
November 6, 2023  
Page Five

**Requests For Production Nos. 3, 21-24.** Similar to its response above regarding Interrogatory No. 1, OpenAI disputes that agreements with Microsoft and other entities regarding products other than Codex and Copilot are relevant to this litigation. Moreover, Plaintiffs' request for "[a]greements with *any* entity regarding *any* products" is overly broad, burdensome, and seeks non-relevant information. Plaintiffs fail to explain why such agreements, earlier iterations of agreements, and communications would be relevant to their allegations, which involve Codex and Copilot. During the meet and confer, you raised Sam Altman's investment into Helion as an example of an agreement you are seeking. Mr. Altman's personal investments are not relevant to this litigation and do not involve OpenAI. There is absolutely no basis for OpenAI to produce such an agreement.

**Requests For Production Nos. 4, 25.** OpenAI has produced publicly available documents regarding the corporate formation of each of the OpenAI defendants. Additional documents regarding the formation of each of the OpenAI defendants are not relevant because Plaintiffs' allegations do not implicate the formation of OpenAI. Please explain why such documents are relevant to any of Plaintiffs' allegations, with reference to specific paragraphs of the Amended Complaint.

Further, OpenAI also states that OpenAI Startup Fund GP I, L.L.C., OpenAI Startup Fund I, L.P., and OpenAI Startup Fund Management, LLC were not involved in the training and development of Codex and/or Copilot. As discussed during the meet and confer, OpenAI requests that Plaintiffs identify the documents that they need from OpenAI to allow these entities to be dismissed from this litigation.

**Request For Production No. 8.** OpenAI agrees that it will make available for inspection relevant Codex models, through an agreed-upon process, and additional documents sufficient to show the operation of Codex.

**Requests For Production Nos. 9, 15-16, 37-39, 44, 46-49.** Plaintiffs mischaracterize the meet and confer between OpenAI and Plaintiffs. The parties did not discuss "sampling" and it is unclear what Plaintiffs mean by that—e.g., what specifically is it that Plaintiffs want OpenAI to "sample" and how? OpenAI suggested that the parties could work on a protocol involving querying, in which Plaintiffs provide a certain number of specific strings of code and/or file names and OpenAI could further investigate how to run queries and provide information. With respect to these requests, OpenAI also pointed out that Plaintiffs have not given Defendants the code that Plaintiffs accuse OpenAI of supposedly ingesting and emitting. As explained during the meet and confer, simply providing repository names is insufficient. Plaintiffs must provide OpenAI with the specific code they are requesting OpenAI to investigate. Once Plaintiffs have done so, OpenAI is willing to revisit these requests. Regarding Plaintiffs' request for documents that describe the contents of the training datasets, OpenAI's supplemental response to Interrogatory No. 13 will provide the requested information.

## |||HARRISON FOERSTER

Joseph Saveri  
November 6, 2023  
Page Six

**Request for Production No. 10.** During the meet and confers, OpenAI indicated that it is willing to investigate Plaintiffs' use of Codex, but stated that it needed to know which Plaintiffs used the Codex API. As reflected in your letter, Plaintiffs agreed to identify the username and account information for any Plaintiffs, if this information exists. We await this information from Plaintiffs.

**Requests for Production Nos. 11-14.** As stated in our response and during the meet and confer, OpenAI is willing to make relevant Codex models available for inspection pursuant to the Amended Protective Order, which will reflect changes to Codex models related to "emission or filtering of output." Communications regarding changes to particular technical functions would be duplicative and cumulative of the changes noted in the source code, and Plaintiffs have not shown why such discovery would be proportional to the needs of this case. Nonetheless, OpenAI is willing to investigate whether such documents exist and further assess whether they are relevant and appropriate for production. To the extent Plaintiffs seek the same information as it relates to Copilot, OpenAI states that it is not aware of changes made to Copilot related to its emission or filtering of output, as Copilot is not a product offered by OpenAI.

**Requests for Production Nos. 17-18.** OpenAI is investigating whether it is possible to provide "prompt-output pairs" for Plaintiffs' code and will update Plaintiffs once that investigation is complete. The parties are to further meet and confer regarding reasonable methodologies to query such data sources to the extent they exist. As requested for Requests for Production Nos. 9, 15-16, 37-39, 44, 46-49, OpenAI asks that Plaintiffs provide Plaintiffs' code they want OpenAI to investigate.

**Requests for Production Nos. 19-20.** Plaintiffs' request is unreasonable. OpenAI does not exercise control over GitHub, and accordingly, cannot "take steps to ensure GitHub produces" documents responsive to these requests.

**Requests for Production Nos. 26-27.** OpenAI has produced publicly available documents regarding the corporate formation of each of the OpenAI defendants, which sufficiently identify the owners of each OpenAI defendant. It is unclear what additional documents in connection with these requests would be relevant to Plaintiffs' allegations in this case.

**Requests for Production Nos. 28-32.** As explained during the meet and confers, OpenAI is open to narrowing these requests, but as written these requests are not a reasonable starting point. Plaintiffs have not explained how *any* documents provided to any US legislator, US regulatory body, the DOJ, the FTC, the SEC, or any non-US regulatory body are relevant to the claims and defenses in this litigation, or proportional to the needs of this case. OpenAI is surprised by Plaintiffs' sudden declaration that the parties are at an impasse when during the meet and confers, Plaintiffs represented that they would come back to OpenAI regarding potential narrowing. To the extent Plaintiffs are no longer interested in discussing a



## MORRISON FOERSTER

Joseph Saveri  
November 6, 2023  
Page Seven

reasonable scope for these requests, please provide an explanation as to how each of these requests and their subparts are relevant to the allegations in this litigation.

**Requests for Production Nos. 33.** The First Amendment protects OpenAI's right to petition the government in connection with generative AI and its business operations. Plaintiffs have not explained how OpenAI's lobbying efforts are relevant to their claims or to determining OpenAI's liability or damages. Please provide an explanation.

**Request for Production No. 34.** OpenAI will produce documents sufficient to show any revenue and costs for Codex between 2021 to present as soon as reasonably practicable.

**Request for Production No. 35.** Although the requested information should be in Plaintiffs' possession, OpenAI will investigate agreements between the named Plaintiffs and OpenAI in response to this request, and to the extent they exist, will produce them as soon as reasonably practicable.

**Requests for Production Nos. 40-43, 45.** As indicated above, documents related to Azure do not fall within the scope of the allegations in this litigation. If you believe otherwise, please explain why. Again, OpenAI is surprised by Plaintiffs' sudden declaration that the parties are at an impasse when during the September 7, 2023 meet and confer, Plaintiffs stated that they would be willing to narrow these requests and represented that it was "fine" if there are no unique documents being produced in response to these requests beyond the agreements between OpenAI and Microsoft pertaining to Codex and/or Copilot.

**Request for Production No. 50.** As explained during the parties' meet and confers, this request is overbroad because, as currently phrased, the request appears to seek documents regarding lawsuits that do not involve any of the products at issue and are not relevant to the claims asserted here. The example discussed during the parties' meet and confers was an OpenAI employee involved in a car accident with a claim of indemnity. Clearly, such a lawsuit would not be relevant. Plaintiffs indicated that they would be willing to narrow this request and now seek "production of complaints filed in federal or state court or complaints submitted to any federal or state law enforcement or regulatory agency." This purported narrowing would still encompass the example of an OpenAI employee involved in a car accident. Again, Plaintiffs offer no explanation for how lawsuits that do not involve any of the products at issue are relevant to the claims asserted here. Moreover, the vast majority of what Plaintiffs seek is publicly available and equally accessible to Plaintiffs.

**Requests for Production 51-52.** OpenAI disputes that these requests "call for discoverable information" and that it refused to produce documents "on the basis that most responsive documents are public." To the extent Plaintiffs seek documents regarding specific ads and promotional material that are publicly available, OpenAI is open to investigating these specific ads and promotional material identified by Plaintiffs that are relevant to their claims.



## MORRISON FOERSTER

Joseph Saveri  
November 6, 2023  
Page Eight

To date, Plaintiffs have not identified any OpenAI ads or promotional materials pertaining to Copilot.

**Request for Production No. 53.** As explained during the meet and confers, this request for complete personnel files of OpenAI employees who worked on Codex is highly invasive and seeks information not relevant to this case. Plaintiffs indicated that they were not interested in the entire history and personnel files of OpenAI employees. During the meet and confers, Plaintiffs proposed narrowing the request to documents regarding employees who raised concerns about the use of particular data as training material for Codex. OpenAI looks forward to Plaintiffs' proposed narrowing, and are willing to further meet and confer regarding the proper scope for this request.

**Request for Production No. 54.** OpenAI states that it has already produced public documents sufficient to show its document retention policies and that Plaintiffs have not raised any specific issues regarding the sufficiency of OpenAI's document retention. Accordingly, OpenAI does not believe any further production is warranted.

**Requests for Production 55-56.** As stated in its initial disclosures, OpenAI will produce the insurance agreements referenced in its initial disclosures as soon as reasonably practicable.

**Request for Production 57.** As OpenAI stated during the parties' September 7, 2023 meet and confer, OpenAI does not know how to search for "ALL DOCUMENTS or COMMUNICATIONS related to or mentioning Plaintiffs" without revealing Plaintiffs' names broadly within OpenAI, which Plaintiffs' counsel has prohibited OpenAI from doing. The parties discussed using search terms and custodians to identify responsive documents and will further meet and confer regarding how to do so as discovery progresses.

Sincerely,



Allyson R. Bennett